

ABSTRACT OF THE DISCLOSURE  
"Expanding Dictation Vocabulary"

In the development of an automatic speech recognition (ASR) system, an extensive study of the basic phonetic alphabet is performed to collect information regarding phonology and phonetics of the language or dialect in question (modern standard Arabic or MSA in this case). In addition, terminological and transcriptional problems are identified with respect to the language or dialect in question. Next, based on feature description (rather than symbol shapes), the symbols in the literature are mapped to a single or more recent phonetic alphabet. Lastly, from a maximal set containing all the phonemes, allophones, and transliteration symbols, a reduced set is created with a compact set of phonetic alphabets. Memory consumption is greatly reduced in a computer system by using this compact set of phonetic alphabets